Similarity-based classification in natural language

Helmar Gust & Carla Umbach
{helmar.gust, carla.umbach}@uos.de

Institute of Cognitive Science, University of Osnabrück

Workshop 'concepts & feature spaces', Osnabrück, June 2010

Two ways of classification

(1) a. Anna ist 1,80m groß.
'Anna is 1,80m.'

b. Anna ist eine kluge Studentin mit beachtlichen Mathekenntnissen.
'Anna is a clever student with considerable math skills.'

rule-based / **categorical** classification
requires lexically given categories

(2) a. Anna ist so groß wie Berta.
'Anna is as tall as Berta.'

b. Anna ist so eine Studentin wie Berta.
'Anna is such a student as Berta .'

similarity-based classification
requires (i) an individual known to the hearer
(ii) relevant dimensions of comparison

Equative comparison constructions

(2) a. Anna ist so groß wie Berta. 'Anna is as tall as Berta.'	**scalar**	**adjectival**
b. Anna ist so eine Studentin wie Berta. 'Anna is such a student as Berta .'	**non-scalar**	**nominal**

➔ How to spell out the semantics of **scalar** as well as **non-scalar** equative comparison in a uniform fashion?

Two types of semantics for adjectival comparison

Degree-based analyses (e.g. Bierwisch 1984, Kennedy 1999)
make use of degrees and measure functions

$\mu_{height}: U \rightarrow D$

$[_{Adj}\ tall] = \lambda x.\ \mu_{height}(x) \geq \mathbf{s}$ — **s** cut-off point given by the comparison class

Vague predicate analyses (e.g. Klein 1980)
reject degrees as part of the ontology

$[_{Adj}\ tall] = \lambda x.\ \mathrm{TALL}(x)$

McNally (2010)
make use of similarity in interpreting (relative) adjective ,
→ vagueness

Adjectives vs. nouns: one vs. multiple dimensions

(2) a. Anna ist so groß wie Berta.
'Anna is as tall as Berta.'

- one dimension
- explicit

→ Anna is similar to Berta **with respect to height**

b. Anna ist so eine Studentin wie Berta.
'Anna is such a student as Berta .'

- multiple dimensions
- implicit

→ Anna is similar to Berta **with respect to the relevant students features**

➔ look inside the noun's meaning
– make the dimensions within the noun available for comparison

Nominal dimensions

Nominal dimensions can be made explicit:

(2) A: Anna ist so eine Studentin wie Berta.
'Anna is such a student as Berta .'

B: In welcher Hinsicht?
'Why?'

A: Na ja, sie, studiert dasselbe Fach, ist nur 1 Semester höher, hat genauso gute Noten, ist mindestens so schlau und kriegt auch ein Stipendium.
' Well, ...same subject, only one semester higher, same grades, at least as clever, also receives scholarship ....'

Nominal dimensions need not use proportional scales

subject:	{Math, CogSci, philosophy, physics, ...}	nominal
semester:	positive integers	proportional
average grade:	{A, B, C, D, E, F}	nominal + order
intelligence:	degrees of intelligence	proportional
scholarship:	{+/-}	binary

Hyp 1: Comparison makes use of dimensions

tall	height:	degrees (real numbers)
student	subject:	{MATH, LINGUISTICS, COGSCI, ...}
	average grade:	{A, B, C, D, E, F}
	semester:	positive integers
	talent:	{+, – }
	height:	degrees

➔ **"generalized measure function"**

Hyp 2: Comparison makes use of predicates

"*A is as tall as B*" true iff A and B are both tiny, small, medium, tall, gigantic, ...

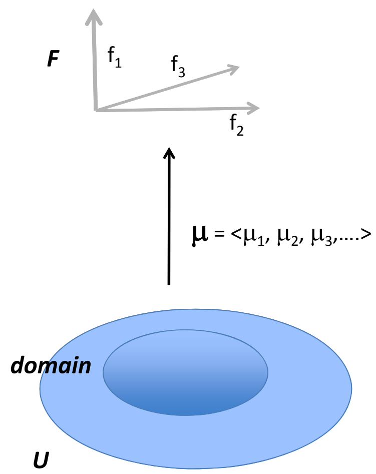
➔ **"generalized measure predicates "**

Extension and "Feature-Representation"

Representation by feature structure
? "Sinn" (Frege)
? concept (Bierwisch)

μ generalized measure function

Extension
domain (μ) $\subset$ universe



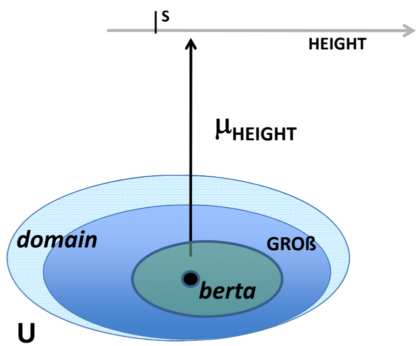
Hyp 1: Comparison makes use of dimensions

One-dimensional case

(to be revised!)

Measure function:

μ_{HEIGHT} (berta)



Hyp 1: Comparison makes use of dimensions (cont.)

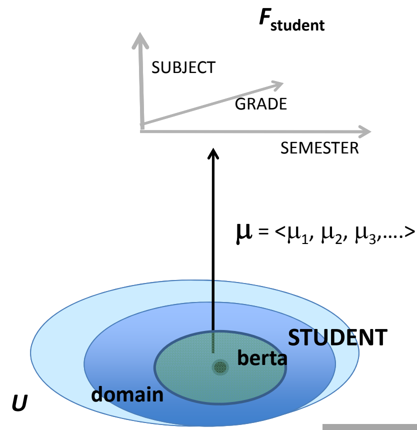
Multi-dimensional case

Generalized measure function

μ_{student} (berta)

$= < \mu_{\text{subject}}, \mu_{\text{semester}}, \mu_{\text{grade}}, \ldots >$ (anne)

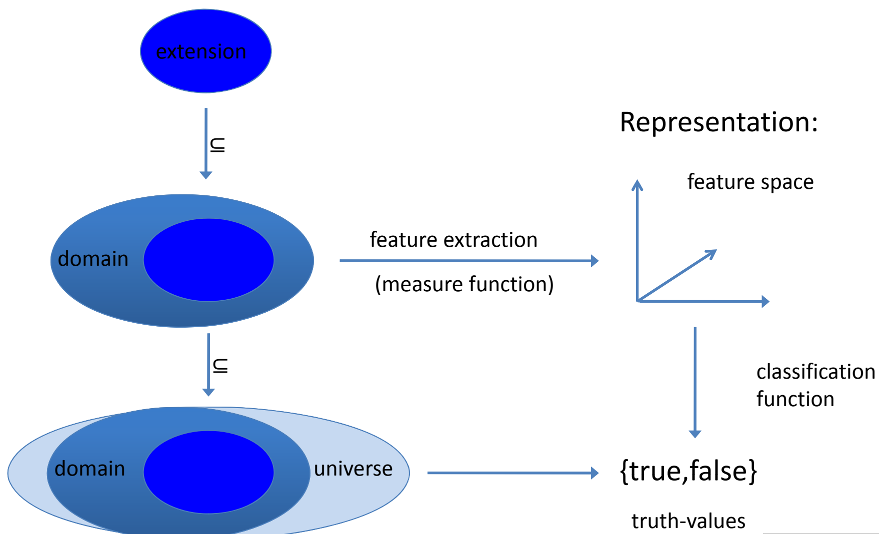
$= \left(\begin{array}{ll} \texttt{SUBJECT:} & \text{MATH} \\ \texttt{SEMESTER:} & 3 \\ \texttt{GRADE:} & \text{A} \end{array} \right)$



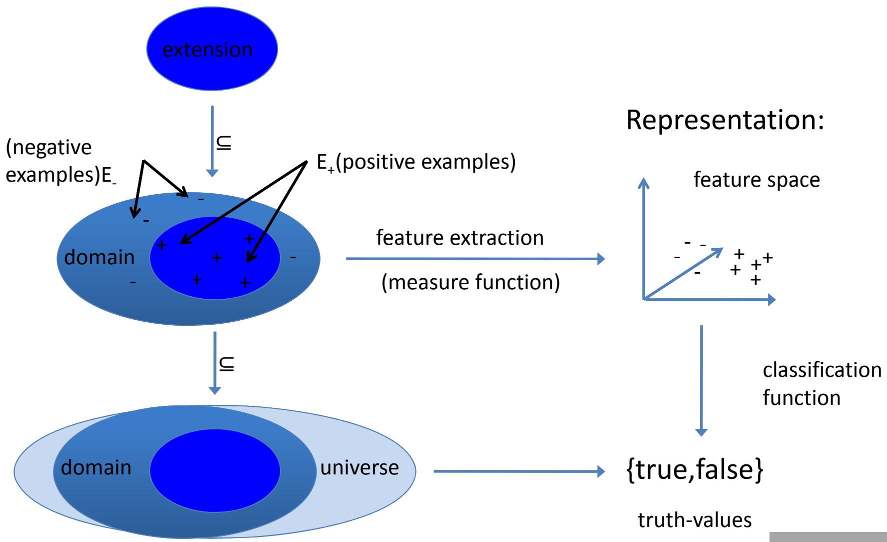
Hyp 2: Comparison makes use of predicates

Generalized Measure Predicates (GMPs)

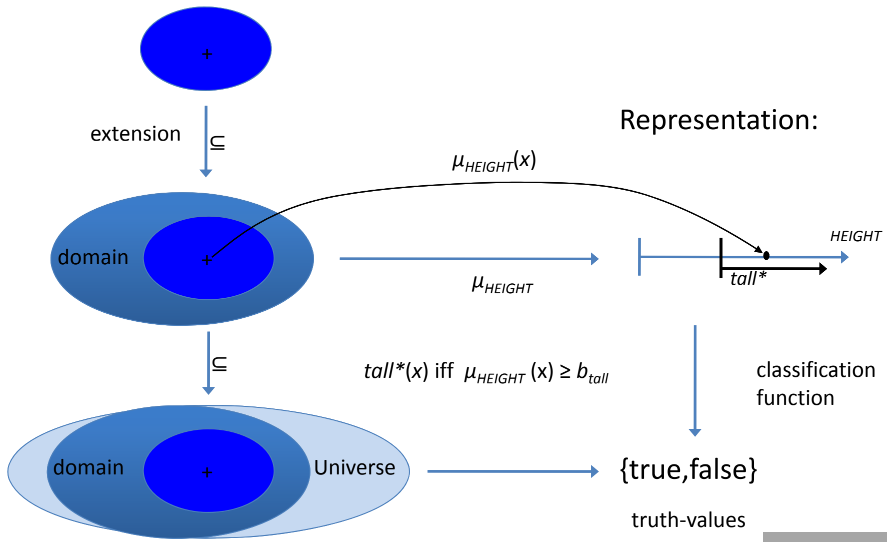
Concepts as predicates



Concepts as predicates



Measure functions (height)



Problem

- Precise values do not matter
- Properties like "*groß*" (tall) are vague
- Even if precise values are mentioned
  - Like "1.80m groß"
  - This doesn't mean exactly 1.80000000m tall
- What matters are the inferences
  - "tall" implies "not small"
  - "tall" implies "taller than medium"
  - etc.

Coherence of Concepts

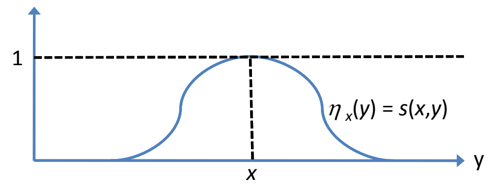
- Continuity
  - slight changes should not change concept membership
- Closedness
  - entities which are 'between' members are members, too
- Stability
  - properties define concepts in a stable way
    - properties do change arbitrarily

Similarity

- Things are similar
  - if they are "near by" in representation space
  - Best of all cases:
    - metrical spaces $d : D \times D \to R^+$
    - similarity measures $s : D \times D \to [0,1]$
  - Similarity measure and metrics are strongly related
    - We can translate between similarity and distance:
      $s(x,y) = c / (c + d(x,y))$
  - Weaker concepts of "near by" :
    - Fuzzy
    - Topology: Neighborhood Systems

Similarity and Fuzzy Sets

- $s : U \times U \to [0,1]$
- $s(x,x) = 1$
- Relation to fuzzy sets: $\eta_x(y) := s(x,y)$
  can be seen as a fuzzy set



namely a fuzzy neighborhood of *x*

Qualitative Fuzzy Sets

- Standard fuzzy sets do not completely solve the qualitative vs. quantitative problem: The values of the membership function are crisp.

- There is the concept of qualitative fuzzy sets: Such a set *X* is specified by a family of membership functions F_X

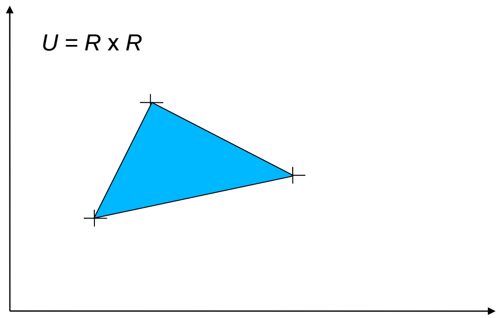
Qualitative Fuzzy Sets

- F_X specifies a qualitative fuzzy set *X* if
  $\forall\ \eta_1, \eta_2 \in F_X$:
  (1) $\eta_1(x) = 1 \leftrightarrow \eta_2(x) = 1$
  (2) $\eta_1(x) = 0 \leftrightarrow \eta_2(x) = 0$
  (3) $\eta_1(x) < \eta_1(y) \leftrightarrow \eta_2(x) < \eta_2(y)$
  (4) $\forall x : d(\eta_1(x), \eta_2(x))$ is small

T. Y. Lin (1997): *NEIGHBORHOOD SYSTEMS: A Qualitative Theory for Fuzzy and Rough Sets.* Advances in Machine Intelligence and Soft Computing, Volume IV. Ed. Paul Wang, 132-155.

Closure Spaces

- *What's a closure?*

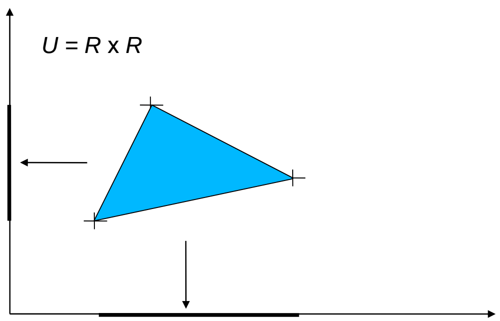


Closure Spaces

- A closure space (U,cl) is a set U together with an operator cl : $\wp(U) \to \wp(U)$ with:
  - $X \subseteq cl(X)$
  - $X \subseteq Y \to cl(X) \subseteq cl(Y)$
  - $cl(cl(X)) = cl(X)$
- In most cases we additionally assume
  - $cl(\{\}) = \{\}$
- X is closed if $X = cl(X)$
  - we define: $closed(X) \leftrightarrow X = cl(X)$
- Classical example:
  - Convex closure in a vector space

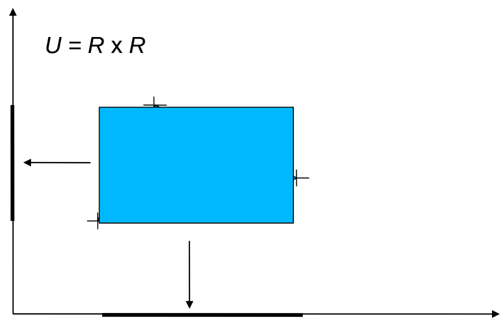
Induced closures

- May be not convex



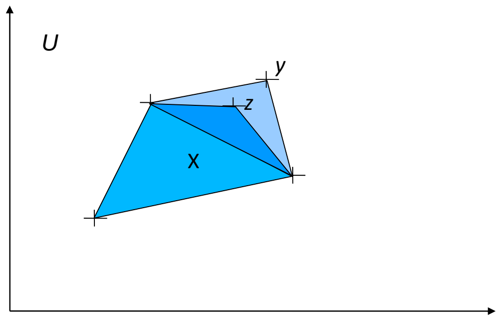
Induced closures

- May be not convex



Convex Closure Spaces

Anti-Exchange Property



$y,z \notin cl(X) \wedge z \in cl(X \cup \{y\}) \rightarrow y \notin cl(X \cup \{z\})$.

Closures in Fuzzy Systems

- Definitions:

$$\alpha\text{-}cut(\eta) := \{x \mid \eta(x) \geq \alpha\}$$

- A fuzzy membership function η over a closure space space (U,cl) is ***closed*** iff

$$\forall \alpha \geq 0 : closed(\alpha\text{-}cut(\eta))$$

- A fuzzy membership function η over a closure space space (U,cl) is ***weakly closed*** iff

$$\forall x,y \in U :$$
$$\eta(x)=1 \rightarrow closed(\eta|_{cl(\{x,y\})})$$

Closures in Fuzzy Systems

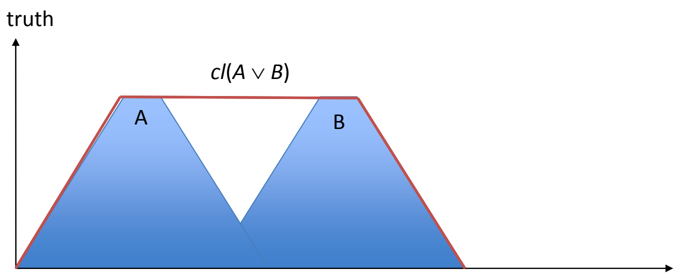
- Remarks:
  - The difference between *closed* and *weakly closed* is relevant only in multi dimensional spaces
  - The core (1-cut) of a weakly closed fuzzy set is always closed
  - Complements of closed sets are in general not closed

Closures in Fuzzy Systems

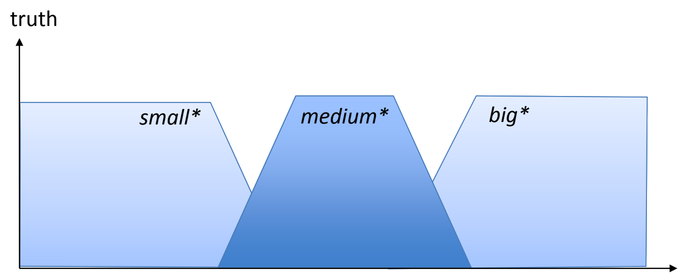
- Hypothesis:

generalized measure predicates are weakly closed qualitative fuzzy sets

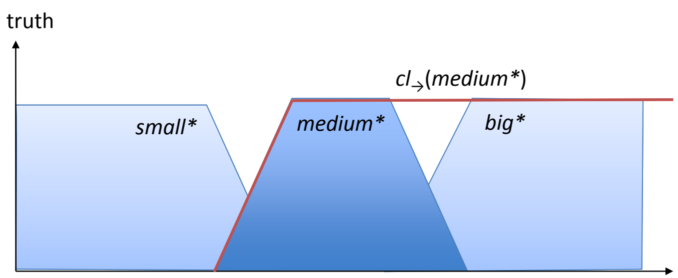
Closures in Fuzzy System



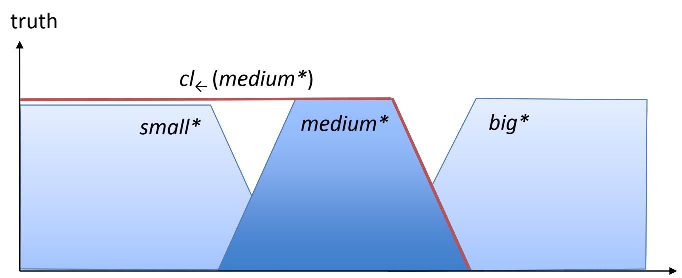
Closures in Fuzzy System



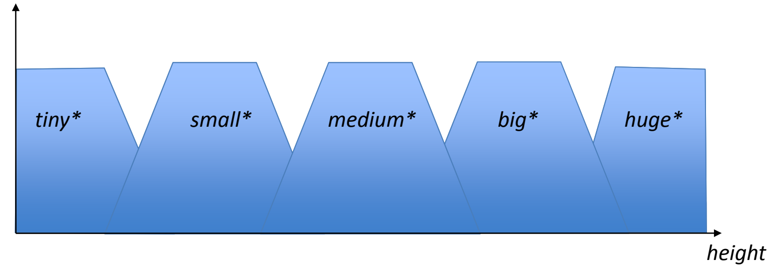
Closures in Fuzzy System



Closures in Fuzzy System



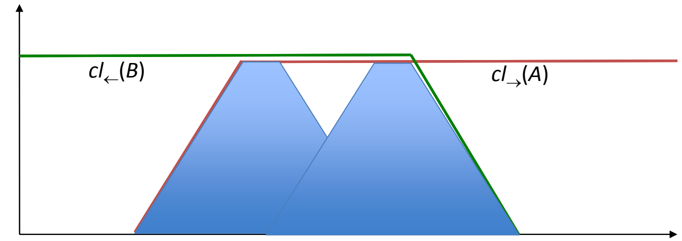
Granularity: Defining a Base B_{height}



- Again: precise values do not matter
- The order matters
- On an ordered dimension we get a partially ordered set of predicates

(Partial) Orders on GMP‘s

- $A < B$ iff $closed(B \backslash A) \wedge closed(A \backslash B) \wedge B \subseteq cl_{\rightarrow}(A)$
- $B > A$ iff $closed(B \backslash A) \wedge closed(A \backslash B) \wedge A \subseteq cl_{\leftarrow}(B)$



Sets of Generalized Measure Predicates

- Given a set of closed base predicates B over a feature F we define a set of predicates P inductively:
  - $B \subseteq P$
  - $A \in P, B \in P \rightarrow A \cap B \in P$
  - $A \in P, B \in P \rightarrow cl(A \cup B) \in P$
- If F is (partially) ordered
  - $A \in P \rightarrow cl_{\rightarrow}(A) \in P$
  - $A \in P \rightarrow cl_{\leftarrow}(A) \in P$

How to use generalized measure functions and generalized measure predicates in the analysis of equative comparison constructions?

(2) a. Anna ist so groß wie Berta. — scalar
'Anna is as tall as Berta.'

$\Rightarrow$ Anna is similar to Berta in height

b. Anna ist so eine Studentin wie Berta. — non- scalar
'Anna is such a student as Berta .'

$\Rightarrow$ Anna is similar to Berta with respect to the relevant student features

The semantics of equative comparison

(2) a. Anna ist so groß wie Berta.
'Anna is as tall as Berta.'

"exactly": $\forall p \in P_{f\text{-height.}}\ p(\mu(A)) \leftrightarrow p(\mu(B))$
"at least": $\forall p \in P_{f\text{-height.}}\ p(\mu(B)) \rightarrow \mathbf{cl}_{\rightarrow}(p)(\mu(A))$

b. Anne ist so eine Studentin wie Berta.
'Anna is such a student as Berta .'

" exactly": $\forall$**relevant** $f_i \in \boldsymbol{F}_{student}\ \forall p \in P_{fi}\ p(\mu(B)) \leftrightarrow p(\mu(A))$

clearly too strong !

Where does similarity come into play?

(i) Similarity as the **basis for the generalized measure predicates** (tiny*, ...)
e.g. instance-based, defined via closed sets and closed membership functions

(ii a) Similarity as a **relation** between individuals **within dimensions**
x and y are similar wrt. dimension f
iff $\forall p \in P_f\ .\ p(f(x)) \leftrightarrow p(f(y))$

(ii b) Similarity as a **relation** between individuals **across dimensions**
x and y are similar wrt. a feature space F
iff $\forall f \in F, \forall p \in P_f\ .\ p(f(x)) \leftrightarrow p(f(y))$

➔ (ii a/b) very strong – weaker notions exist.
e.g., count the number of predicates two items have in common,
weight predicates according to relevance, ...

Conclusion

→ **Framework to implement similarity**
- closure spaces
- qualitative fuzzy sets

→ **Semantic interpretation for similarity-based classification in natural language,** in particular equative comparison
- Generalized measure functions (mappings from individuals into multi-dimensional spaces)
- Generalized measure predicates (predicates over dimensions, defined as weakly closed qualitative fuzzy sets)
- Comparison based on similarity defined via generalized measure predicates

References

Bierwisch, M. (1984), 'The semantics of gradation', in Bierwisch, M. and E. Lang (eds.), *Dimensional Adjectives, pp. 71-262, Springer Verlag.*

Gärdenfors , P. (2000(Conceptual Spaces. MIT Press

Kennedy, C. (1999), *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison. Garland Press, New York.*

Klein, E. (1980), 'The semantics of positive and comparative adjectives', *Linguistics and Philosophy*, 4: 1-45.

T. Y. Lin (1997): *NEIGHBORHOOD SYSTEMS: A Qualitative Theory for Fuzzy and Rough Sets.* Advances in Machine Intelligence and Soft Computing, Volume IV. Ed. Paul Wang, 132-155.

McNally , Louise (2010) The Relative Role of Property Type and Scale Structure in Explaining the Behavior of Gradable Adjectives. Proceedings of the ESSLLI 2009 workshop on Vagueness in Communication.

Tversky, Amos (1977) Features of Similarity. Psychological Review Vol. 84, (327-352)

Pawlak, Zdzisław (1982). "Rough sets". *International Journal of Parallel Programming* 11 (5): 341–356.

Pawlak, Zdzisław (1991). *Rough Sets: Theoretical Aspects of Reasoning About Data*. Dordrecht: Kluwer Academic Publishing.